So far, detection under:

- Neyman-Pearson criteria (max P_D s.t. P_{FA} = constant): likelihood ratio test, threshold set by P_{FA}
- minimize Bayesian risk (assign costs to decisions, have priors of the different hypotheses): likelihood ratio test, threshold set by priors+costs
    - minimum probability of error = maximum a posteriori detection
        - maximum likelihood detection = minimum probability of error with equal priors
- known deterministic signals in Gaussian noise: correlators
- random signals: estimator-correlators, energy detectors

All assume knowledge of $p(x; \mathcal{H}_0), p(x; \mathcal{H}_1)$

Motivation

- What if don't know the distribution of x under the two hypotheses?
- What if under hypothesis 0, distribution is in some set, and under hypothesis 1, this distribution lies in another set - can we distinguish between these two?

Composite hypothesis testing

We now ask the question: "how can we detect a signal presence when the probability density functions under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ are not exactly known?" We do assume that the pdfs are parametrized and that these parameters may or may not be known. More specifically, under $\mathcal{H}_0$ the set of unknown parameters is $\theta_{\mathbf{0}}$, while under $\mathcal{H}_1$ the set of parameters is $\theta_{\mathbf{1}}$.

Composite hypothesis testing summary

The pdfs of the received signal $\mathbf{x}$ under the two hypotheses are $p(\mathbf{x}, \theta_{\mathbf{0}}, \mathcal{H}_0)$ and $p(\mathbf{x}, \theta_{\mathbf{1}}, \mathcal{H}_1)$. There are three main approaches one can take:

1. Hope for a Uniformly Most Powerful (UMP) test
2. Bayesian approach
3. Generalized likelihood ratio test

1. Uniformly Most Powerful test (UMP)

The pdfs of the received signal $\mathbf{x}$ under the two hypotheses are $p(\mathbf{x}, \boldsymbol{\theta_0}, \mathcal{H}_0)$ and $p(\mathbf{x}, \boldsymbol{\theta_1}, \mathcal{H}_1)$, where under $\mathcal{H}_0$ the set of unknown parameters is $\boldsymbol{\theta_0}$, while under $\mathcal{H}_1$ the set of parameters is $\boldsymbol{\theta_1}$.

- form the regular likelihood ratio test assuming the parameters are known.
- when the test that yields the highest P_D for a given P_{FA} does **not** depend on the unknown parameters $\boldsymbol{\theta_0}, \boldsymbol{\theta_1}$, this is called a UMP test.
- When a UMP test does not exist you can use a sub-optimal test and compare its performance to the *clairvoyant* detector, i.e. the detector who knows the parameters completely. For a UMP to exist, you must be able to frame the test as a "one-sided parameter test", i.e. under $\mathcal{H}_0$ your unknown parameter is some value, while under $\mathcal{H}_1$ the test for the parameter is *one-sided* (e.g. $A > 0$ or $A < 0$ but not $A \neq 0$.)

Example: DC level in WGN with unknown amplitude A>0

$$\begin{aligned} \mathcal{H}_0 &: x[n] = w[n], \qquad n = 0, 1, \cdots N-1 \\ \mathcal{H}_1 &: x[n] = A + w[n], \qquad w[n] \sim \mathcal{N}(0, \sigma^2) \end{aligned}$$

where A is unknown, except that we know that $A > 0$. How can we distinguish between $\mathcal{H}_0$ and $\mathcal{H}_1$?

- What is the test with perfect knowledge of A?
- What is the test without knowledge of A?
- How does the performance of the two compare?

An alternative UMP view of the previous example

- Can you translate the previous example into a test on A?

- What happens when you don't know whether A>0 or A<0?

- What happens to the performance?

2. Bayesian approach

The pdfs of the received signal $\mathbf{x}$ under the two hypotheses are $p(\mathbf{x}, \boldsymbol{\theta_0}, \mathcal{H}_0)$ and $p(\mathbf{x}, \boldsymbol{\theta_1}, \mathcal{H}_1)$, where under $\mathcal{H}_0$ the set of unknown parameters is $\boldsymbol{\theta_0}$, while under $\mathcal{H}_1$ the set of parameters is $\boldsymbol{\theta_1}$.

Assuming we have reasonable prior probabilities on the parameters $\boldsymbol{\theta_0}, \boldsymbol{\theta_1}$ given by $p(\boldsymbol{\theta_0})$ and $p(\boldsymbol{\theta_1})$ we can integrate out the dependence of the the pdfs on the unknown parameters and for the Bayesian likelihood ratio:

$$\frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} = \frac{\int p(\mathbf{x}|\boldsymbol{\theta_1}; \mathcal{H}_1) p(\boldsymbol{\theta_1}) d\boldsymbol{\theta_1}}{\int p(\mathbf{x}|\boldsymbol{\theta_0}; \mathcal{H}_0) p(\boldsymbol{\theta_0}) d\boldsymbol{\theta_0}} > \gamma'$$

The choice of prior distributions and the multi-variate integration (often no nice closed form) can make this a tough approach.

Example: DC level in AWGN with Gaussian prior

$$\mathcal{H}_0 : x[n] = w[n], \qquad n = 0, 1, \cdots N-1$$
$$\mathcal{H}_1 : x[n] = A + w[n], \qquad w[n] \sim \mathcal{N}(0, \sigma^2), \;\; A \sim \mathcal{N}(0, \sigma_A^2)$$

and A and $w[n]$ are independent.

- What is the test now?

3. Generalized likelihood ratio test (GLRT)

when $\boldsymbol{\theta_0}, \boldsymbol{\theta_1}$ are unknown we can form a detector based on the Maximum Likelihood Estimates (MLE) of the parameters based on the given sequence. Calling the MLE $\hat{\boldsymbol{\theta_0}} = \arg\max p(\mathbf{x}; \boldsymbol{\theta_0}, \mathcal{H}_0)$ and $\hat{\boldsymbol{\theta_1}} = \arg\max p(\mathbf{x}; \boldsymbol{\theta_1}, \mathcal{H}_1)$, we can form the generalized likelihood ratio $L_G(\mathbf{x})$ and all its equivalent version:

$$\begin{aligned} L_G(\mathbf{x}) &= \frac{p(\mathbf{x}; \hat{\boldsymbol{\theta_1}}, \mathcal{H}_1)}{p(\mathbf{x}; \hat{\boldsymbol{\theta_0}}, \mathcal{H}_0)} \\ &= \frac{\max_{\theta_1} p(\mathbf{x}; \boldsymbol{\theta_1}, \mathcal{H}_1)}{\max_{\theta_0} p(\mathbf{x}; \boldsymbol{\theta_0}, \mathcal{H}_0)} \end{aligned}$$

Example: GLRT of DC level in AWGN

$$\mathcal{H}_0 : x[n] = w[n], \qquad n = 0, 1, \cdots N-1$$
$$\mathcal{H}_1 : x[n] = A + w[n], \qquad w[n] \sim \mathcal{N}(0, \sigma^2)$$

where A is unknown. How can we distinguish between $\mathcal{H}_0$ and $\mathcal{H}_1$?

Example: GLRT of DC level in WGN with 2 unknown parameters

$$\mathcal{H}_0 : x[n] = w[n], \qquad n = 0, 1, \cdots N-1$$
$$\mathcal{H}_1 : x[n] = A + w[n], \qquad w[n] \sim \mathcal{N}(0, \sigma^2)$$

where A and σ^2 are both unknown. How can we distinguish between $\mathcal{H}_0$ and $\mathcal{H}_1$?

Performance

The idea behind getting the performance (probability of detection) is to determine the distribution of the test statistic $T(\mathbf{x})$, from which you can obtain the $\Pr\{T(\mathbf{x}) > \gamma'\}$ to obtain the performance.

While this is not always analytically easy/possible to do, conceptually and computationally, this may be usually be accomplished.

Alternatives to the GLRT

The generalized likelihood ratio test (GLRT) requires determining maximum likelihood estimates of the unknown parameters $\theta_{\mathbf{0}}$ and $\theta_{\mathbf{1}}$. This may or may not be easy to do. We thus look at two different tests whose asymptotic (as the number of samples $N \to \infty$) performance are the same as that of the GLRT. The reason we're interested in these tests is that they may be much easier to compute than the GLRT, depending on the scenario.

- the Wald test
- the Rao test

Nuisance parameters are parameters in the pdfs which may be unknown but are the same under different hypotheses. They are a nuisance as you still need to consider them but they don't necessarily help you discriminate between the different hypotheses. The Wald and Rao tests take on a more generalized form when nuisance parameters are present (see page 214-215), we present their simpler form when no nuisance parameters are present.

Asymptotically equivalent tests

Assume no nuisance parameters and that:

$$\mathcal{H}_0 : \theta = \theta_{\mathbf{0}}$$
$$\mathcal{H}_1 : \theta \neq \theta_{\mathbf{0}}$$

The following three tests are asymptotically equivalent as $N \to \infty$

- Generalized likelihood ratio test:

$$L_G(\mathbf{x}) = \frac{p(\mathbf{x}; \hat{\theta_{\mathbf{1}}}, \mathcal{H}_1)}{p(\mathbf{x}; \hat{\theta_{\mathbf{0}}}, \mathcal{H}_0)}, \text{ where } \theta_{\mathbf{0}}, \theta_{\mathbf{1}} \text{ are the MLE of } \theta_{\mathbf{0}}, \theta_{\mathbf{1}} \text{ under } \mathcal{H}_0, \mathcal{H}_1$$

- Wald test with test statistic $T_W(\mathbf{x})$ given by

$$T_w(\mathbf{x}) = \left(\hat{\theta_{\mathbf{1}}} - \theta_{\mathbf{0}}\right)^T I(\hat{\theta_{\mathbf{1}}}) \left(\hat{\theta_{\mathbf{1}}} - \theta_{\mathbf{0}}\right), \text{ where } \hat{\theta_{\mathbf{1}}} \text{ is the MLE of } \theta_{\mathbf{1}},$$

  and $I(\theta)$ is the Fisher information matrix with components

$$[I(\theta)]_{ij} = -E\left[\frac{\partial^2 \ln p(\mathbf{x}; \theta)}{\partial \theta_i \theta_j}\right] \text{ evaluated at } \theta.$$

  Here we assume $\theta = [\theta_1 \; \theta_2; \cdots \theta_M]^T$, for M unknown parameters.

- Rao test with test statistic $T_R(\mathbf{x})$:

$$T_r(\mathbf{x}) = \left.\frac{\partial \ln p(\mathbf{x}; \theta)}{\partial \theta}\right|_{\theta=\theta_{\mathbf{0}}}^T I^{-1}(\theta_{\mathbf{0}}) \left.\frac{\partial \ln p(\mathbf{x}; \theta)}{\partial \theta}\right|_{\theta=\theta_{\mathbf{0}}}$$

Example: DC level in AWGN

- What are the GLRT, Wald and Rao tests for detecting an unknown DC level A in WGN? Assume A does not equal 0, but otherwise unknown.

Example: DC level in non-Gaussian noise

Suppose we wish to detect an unknown DC level in iid NON-GAUSSIAN noise,

$$\mathcal{H}_0 : x[n] = w[n]$$
$$\mathcal{H}_1 : x[n] = A + w[n]$$

where A is an unknown constant $-\infty < A < \infty$ and $\{w[0], w[1], \cdots, w[n]\}$ are IID with non-Gaussian noise with pdf

$$p(w[n] = \frac{1}{a\sigma}\Gamma(\frac{5}{4})2^{\frac{5}{4}} \exp\left[-\frac{1}{2}\left(\frac{w[n]}{a\sigma}\right)^4\right], \quad -\infty < w[n] < \infty,$$

where $a = 1.4464$. Find detectors - turns out the Rao test is the only really feasible one.